Ann. Inst. Statist. Math.
Vol. 50, No. 1, 187–202 (1998)

A REMARK ON A FOURIER BOUNDING METHOD OF PROOF FOR CONVERGENCE OF SUMS OF PERIODOGRAMS*

A. G. BENN AND R. J. KULPERGER

Department of Statistical and Actuarial Sciences,
The University of Western Ontario, London, Canada N6A 5B7

(Received June 25, 1996; revised May 6, 1997)

Abstract. This paper studies sums of periodograms in a random field setting. In a one dimensional or time series setting these can be studied using a method of cumulants, as done by Brillinger. This method does not carry over well to the random field case. Instead one should apply an argument as used by Rosenblatt. In order to have asymptotically correct confidence intervals, one needs to center these sums properly in the random field case.

Key words and phrases: Finite Fourier transforms, periodograms, random field, cumulants.

1. Introduction

Brillinger (1981) in his time series book, and Rosenblatt (1985) have studied sums of periodograms. Brillinger's method for a real time series is based on cumulant calculations of finite Fourier transforms and the calculus of Leonov and Shiryaev (1959) applied to second degree polynomials of the finite Fourier transforms. The observed data sample size is T. These are very nice as many complicated calculations are easily bounded using spectral density approximations with a $O(T^{-1})$ error bound. Rosenblatt on the other hand studied real time series, as well as spatial time series, but approximated sums of periodograms by a quadratic form in the original real time or spatial domain. The computations are not quite as clean as the more restrictive Brillinger Fourier transform 1D cumulant calculations.

One problem that is addressed in this paper is the applicability of Brillinger's cumulant method to a spatial process, that is a 2D random field. See Section 2 for the example, where it is shown that the nice Fourier bounding method does not work. The process is observed on a $T \times T$ lattice. Basically the failure of the bounding method is due to the error term being of size $O(T^{-1})$, and the rate of

* This work has been supported by the Natural Sciences and Engineering Research Council of Canada.

convergence for the sum of the periodograms being $\sqrt{T^2} = T$, which is of the same size. The big O term cannot be reduced to a little o.

A second issue involves a small but important refinement of Rosenblatt's result for sums of covariances in Z^2 (Rosenblatt (1985)), coupled with a standard time series trick of modifying the divisor or the number of terms in the numerator of a lag covariance in time series, such as given in Brockwell and Davis ((1991), Chapter 1). This is useful for being able to use all the data for computing a sample estimate. In the time series setting one can use

$$\hat{c}(h) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(X_{i+h} - \bar{X})$$

as the sample covariance to study the distribution of the sample covariance. However in application one would typically use, for $h \geq 0$,

$$\tilde{c}(h) = \frac{1}{n} \sum_{i=1}^{n-h} (X_i - \bar{X})(X_{i+h} - \bar{X})$$

as the sample covariance. Asymptotically both are equivalent in a one dimensional time series.

In a spatial setting, without the proper centering, a limiting normal distribution is still obtained, but one with non-zero mean. This would lead to incorrect asymptotic confidence intervals. This problem is addressed in Section 4. This problem does not come up with a 1D time series, since the size of the bias tends to zero. In 2D however the bias does not converge to zero, so the proper centering is needed. A similar result is discussed in Guyon (1982). The proper centering discussed here deals with centering the biased empirical covariances properly. Guyon (1982) deals with the problem by replacing the biased empirical covariances with the unbiased covariance estimates. It should be noted that our solution allows one to still make use of fast Fourier transform algorithm, while Guyon's solution does not allow one to take advantage of the fast Fourier transforms. The use the fast transforms is important in an image analysis problem with a typically large sample size of T^2 with $T = 512$.

2. Two-D example of Fourier and cumulant method

This section will give a summary of the Brillinger 1D proof, and the 2D MA(2) example; actually an overview. The computational details are given in a subsection. First a review of the 1D Brillinger proof is given.

2.1 *The 1D Fourier method*

Brillinger (1981, section 5.10) studies the estimation of the spectral measure, obtaining the weak convergence to a Gaussian process. In the proof of his Theorem 5.10.1, he obtains the L-th cumulant of the normalized spectral measure, arguing that it is $O(T^{(1-L/2)})$, which tends to 0 for $L \geq 3$. The notation cum_L means an L-th cumulant.

This term comes specifically from approximating a joint cumulant of

$$d^{(T)}(\lambda) = \sum_{t=0}^{T-1} e^{-i\lambda t} X_t$$

as

$$\begin{aligned}
&\operatorname{cum}_L(d^T(\lambda_1), \ldots, d^T(\lambda_L)) \\
&= \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} e^{-i\lambda_1 t_1} \cdots e^{-i\lambda_L t_L} \operatorname{cum}_L(t_1 - t_L, \ldots, t_{L-1} - t_L) \\
&= \sum_{t_L}^{T-1} \sum_{u_1} \cdots \sum_{u_{L-1}} e^{-i(\lambda_1 u_1 + \cdots + \lambda_{L-1} u_{L-1})} \operatorname{cum}_L(u_1, \ldots, u_{L-1}) e^{-i(\lambda_1 + \lambda_L) t_L} \\
&= \Delta^{(T)}(\lambda_1 + \cdots + \lambda_L)(2\pi)^{L-1} f_L(\lambda_1, \ldots, \lambda_{L-1}) + O(T),
\end{aligned}$$

where $u_j = t_j - t_L$, and the function $\Delta^{(T)}$ is given by

$$\Delta^{(T)}(\lambda) = \sum_{t=0}^{T-1} e^{-i\lambda t}. \tag{2.1}$$

At the special frequencies of the form $\lambda_{j,T} = \frac{2\pi j}{T}$, $\Delta^{(T)}(\lambda_{j,T})$ equals T if $j = 0 (\bmod T)$, and 0 otherwise. This property of $\Delta^{(T)}$ is a consequence of the oscillatory nature of the complex exponential, that is a Fourier property. The $O(T)$ term comes from bounding a sum of the form

$$\left| \sum_{t=0}^{T-1} e^{-i\lambda t} \right| \le \sum_{t=0}^{T-1} |e^{-i\lambda t}| = T.$$

That is at an appropriate stage one bounds some complex exponentials by 1.

As an immediate consequence one obtains the cumulants of the normalized finite Fourier transforms

$$\begin{aligned}
&\operatorname{cum}_L(T^{-1/2} d^{(T)}(\lambda_1), \ldots, T^{-1/2} d^{(T)}(\lambda_L)) \\
&\qquad = T^{-L/2} \Delta^{(T)}(\lambda_1 + \cdots + \lambda_L) f_L(\lambda_1, \ldots, \lambda_L) + O(T^{(1-L/2)}).
\end{aligned} \tag{2.2}$$

Using (2.2) and the Leonov-Shiryaev calculus applied to periodograms $I^{(T)}(\lambda) = (2\pi)^{-1} d^{(T)}(\lambda) d^{(T)}(-\lambda)$, one can obtain the asymptotic normality of the spectral measure, and similarly for weighted sums or periodograms.

For simplicity take the weights to be 1, that is consider

$$S_T = \sum_{t=1}^{T-1} I^{(T)}(\lambda_{t,T}) \tag{2.3}$$

where $\lambda_{t,T} = 2\pi t/T$. Then an L-th cumulant of S_T is

$$\text{(2.4)} \qquad \begin{aligned} \text{cum}_L(S_T) &= \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} \text{cum}(I^{(T)}(\lambda_{t_1,T}), \ldots, I^{(T)}(\lambda_{t_L,T})) \\ &= \sum_{\mathcal{P}} \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} \text{cum}(P_1) \cdots \text{cum}(P_h) \end{aligned}$$

where the outer sum in the last term is over all indecomposable partitions (Brillinger (1981)) of the table

$$\begin{matrix} t_1 & -t_1 \\ \vdots & \\ t_L & -t_L. \end{matrix}$$

The terms $\text{cum}(P_j)$ are the joint cumulants of the discrete or finite Fourier transforms $T^{-1/2}d^{(T)}(\lambda_{t,T})$ for the t in the partition set P_j in the indecomposable partition.

For example, one of the indecomposable partitions is the case where there are L partition sets each of size 2. For this partition set, the contribution to the sum (2.4) is

$$\begin{aligned} &\sum_{t_1} \cdots \sum_{t_L} \text{cum}_2(P_1) \cdots \text{cum}_2(P_L) \\ &\qquad = T^{-L} \sum_{t_1} \cdots \sum_{t_L} \{(\Delta^{(T)} + O(T)) \cdots (\Delta^{(T)} + O(T))\} \\ &\qquad = T^{-L} T O(T^L) = O(T). \end{aligned}$$

Therefore

$$\text{cum}_L\{T^{-1/2}(S_T - \text{E}(S_T))\} = O(T^{1-L/2})$$

which tends to 0 for $L \geq 3$, where S_T is given by (2.3). The second cumulant can be computed directly, thus yielding a normal limit for S_t.

This proof is very nice, and simple, using to advantage Fourier approximations to order $O(T)$ for cumulants of the finite Fourier transforms. Can one use such a proof in a random field case by perhaps a more more careful use of the term giving rise to $O(T)$?

2.2 *The 2D $MA(1)$ example*

Let $\{\epsilon_{i,j} : i, j \in Z^2\}$ be an iid mean 0 array, with all moments finite. For a given set of non-zero numbers $\alpha_{a,b}$, $a, b = -1, 0, 1$, let $X_{i,j} = \sum_{a,b} \alpha_{i-a,j-b}\epsilon_{a,b}$. Then $\{X_{i,j} : i, j \in Z^2\}$ is a 1 dependent moving average process in Z^2. This is the example for this section. The dependence parameter is the matrix

$$A = \begin{pmatrix} \alpha_{1,-1} & \alpha_{1,0} & \alpha_{1,1} \\ \alpha_{0,-1} & \alpha_{0,0} & \alpha_{0,1} \\ \alpha_{-1,-1} & \alpha_{-1,0} & \alpha_{-1,1} \end{pmatrix}.$$

This process satisfies the mixing conditions of Brillinger (1981). For this process, X_{i_1,j_1} and X_{i_2,j_2} are independent if $|i_1 - i_2| > 2$ or $|j_1 - j_2| > 2$. We can even simplify the dependence structure by taking

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This results in X_{i_1,j_1} and X_{i_2,j_2} being independent for $|i_1 - i_2| > 1$ or $|j_1 - j_2| > 1$, that is X is a 1 dependent process. For the rest of this section we will use a 1 dependent process indexed by Z^2. The process above demonstrates the existence of such a process.

The observed data from the process X is $\{X_{t_1,t_2} : 0 \le t_1, t_2 \le T - 1\}$. That is the observation is on a $T \times T$ rectangle, and the asymptotics are as $T \to \infty$. The finite Fourier transforms are

$$d^{(T)}(\lambda) = \sum_{t=0}^{T-1} \exp\{-i\langle \lambda, t\rangle\} X_t,$$

where the sum is over $t = (t_1, t_2)$, $0 \le t_1, t_2 \le T - 1$, and the argument is $\lambda = (\lambda_1, \lambda_2)$. It should be clear from the context when an argument is a vector. Later as needed some of the vector arguments will be evaluated more explicitly.

The periodograms are given by $I^{(T)}(\lambda) = |d^{(T)}(\lambda)|^2/(T^2)$. Note that usually there is a factor of $(2\pi)^2$ in the denominator, but for notational convenience we omit it. This is because the primary concern of this note is to study the usefulness of the Fourier method of Brillinger (1981), as described earlier, in the random field case.

In a completely analogous manner to (2.2), one obtains

$$\begin{aligned} (2.5)\quad & \mathrm{cum}_L\{T^{-1}d^{(T)}(\lambda_1), \ldots, T^{-1}d^{(T)}(\lambda_L)\} \\ &= T^{-L} \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} e^{-i\langle \lambda_1, t_1\rangle} \cdots e^{-i\langle \lambda_L, t_L\rangle} \mathrm{cum}_L\{X_{t_1}, \ldots, X_{t_L}\} \\ &= T^{-L}\{\Delta_{T,2}(T\{\lambda_1 + \cdots + \lambda_L\}) f_L(\lambda_1, \ldots, \lambda_{L-1}) + O(T^{-1})\} \end{aligned}$$

where $\Delta_{T,2}(\lambda) = \Delta_{T,2}((\gamma_1, \gamma_2)) = \Delta^{(T)}(\gamma_1)\Delta_T(\gamma_2)$ at argument $\lambda = (\gamma_1, \gamma_2)$, and where $\Delta^{(T)}$ is given by (2.1). The $O(T^{-1})$ comes about for the same reason as before, that is by bounding the complex exponential terms by 1.

The periodogram sum is now of the form

$$S_T = \sum_{t=1}^{T-1} I^{(T)}(\lambda_{t,T}),$$

where the sum is over $t = (t_1, t_2)$, $1 \le t_a \le T - 1$, and $\lambda_{t,T} = 2\pi \frac{(t_1,t_2)}{T}$.

By the Leonov and Shiryaev calculus of cumulants, one obtains the analogue of (2.4).

$$\text{cum}_L(S_T) = \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} \text{cum}(I^{(T)}(\lambda_{t_1,T}) \cdots I^{(T)}(\lambda_{t_L,T})) \tag{2.6}$$
$$= \sum_{\mathcal{P}} \sum_{t_1}^{T-1} \cdots \sum_{t_L}^{T-1} \text{cum}(P_1) \cdots \text{cum}(P_h)$$

where the sums are now over $t_j = (t_{j,1}, t_{j,2})$, $1 \leq t_{j,a} \leq T-1$. Consider an arbitrary partition. The number of partition sets is h. Only partition sets of size 2 or greater need to be considered, as the cumulant of partition set of size 1 equals 0. Thus the maximum of h is L. If any partition set is of size greater than 2, then the number of partition sets is $h < L$. In this case the L-th cumulant contribution to (2.6) is $O(T^{-L+h}) \to 0$. If all the partition sets are of size 2, then the cumulant contribution to (2.6) using (2.5) is $O(T^{-L+h}) = O(T^0) = O(1)$ which may or may not tend to zero. A more precise bound or calculation is needed for the case of partition sets of size 2. The question is can this be done to show asymptotic normality? The answer is no. It is shown by an explicit calculation for the third cumulant case $L = 3$ in Section 3 that the third cumulant does tend to zero. Note that this is a tedious calculation, and it is only shown for the third cumulant.

In the rest of this section, we now consider a more careful bound of the partition sets of size 2 for the third cumulant. Consider (2.5) for $L = 2$ and write explicitly the $O(T^{-1})$ remainder term as $R^{(T)}(\lambda_1, \lambda_2)$.

For a 1 dependent process, it can be shown that

$$\begin{aligned}\text{cum}_2(&d^{(T)}(\lambda^{(1)}), d^{(T)}(\lambda^{(2)})) \\ &= (2\pi)^2 \Delta^{(T)}(\lambda^{(1)} + \lambda^{(2)}) f(\lambda^{(1)}) + R^{(T)}(\lambda^{(1)}, \lambda^{(2)})\end{aligned} \tag{2.7}$$

where $R^{(T)}$ is the remainder, I is an indicator function and

$$\begin{aligned}R^{(T)}&(\lambda^{(1)}, \lambda^{(2)}) \\ &= \sum_{t=0}^{T-1} \sum_{u=-\infty}^{\infty} \{(I(-t_1 \leq u_1 \leq T - t_1 - 1) I(-t_2 \leq u_2 \leq T - t_2 - 1) - 1) \\ &\qquad \cdot (I(-1 \leq u_1 \leq 1) I(-1 \leq u_2 \leq 1)) \\ &\qquad \cdot e^{-i\langle t, (\lambda^{(1)} + \lambda^{(2)})\rangle} e^{-i\langle \lambda^{(1)}, u\rangle} c_2(u)\}.\end{aligned}$$

Then,

$$\begin{aligned}|R^{(T)}&(\lambda^{(1)}, \lambda^{(2)})| \\ &\leq \sum_{t=0}^{T-1} \sum_{u=-\infty}^{\infty} |c_2(u)| |I(-1 \leq u_1 \leq 1) I(-1 \leq u_2 \leq 1)| \\ &\quad \cdot |I(-t_1 \leq u_1 \leq T - t_1 - 1) I(-t_2 \leq u_2 \leq T - t_2 - 1) - 1| \\ &= \sum_{u=-1}^{1} |c_2(u)| \sum_{t=0}^{T-1} |I(-t_1 \leq u_1 \leq T - t_1 - 1) I(-t_2 \leq u_2 \leq T - t_2 - 1) - 1|.\end{aligned}$$

Consider $u = (-1, -1)$ and evaluate the inner sum. The inner sum evaluates here to $2T - 1$. Thus,

$$|R^{(T)}(\lambda^{(1)}, \lambda^{(2)})| \leq O(T).$$

Now,

$$\begin{aligned}\mathrm{cum}_3 \Bigg\{ \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(1)}=1}^{T-1} I^{(T)}(\lambda_{j^{(1)}}), \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(2)}=1}^{T-1} I^{(T)}(\lambda_{j^{(2)}}),& \\ \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(3)}=1}^{T-1} I^{(T)}(\lambda_{j^{(3)}}) \Bigg\}& \\ = \frac{1}{T^{12}} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} \{\mathrm{cum}_3(d^{(T)}(\lambda_{j^{(1)}}) d^{(T)}(-\lambda_{j^{(1)}}),& \\ d^{(T)}(\lambda_{j^{(2)}}) d^{(T)}(-\lambda_{j^{(2)}}),& \\ d^{(T)}(\lambda_{j^{(3)}}) d^{(T)}(-\lambda_{j^{(3)}}))\}.&\end{aligned}$$

The Leonov-Shiryaev calculus of cumulants (see Brillinger (1981)) allows one to calculate the cumulants above, in terms of indecomposable partitions. Consider a particular indecomposable partition consisting of 3 sets or partition elements.

$$\begin{aligned}\frac{1}{T^{12}} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} \{ \mathrm{cum}_2(d^{(T)}(\lambda_{j^{(1)}}), d^{(T)}(-\lambda_{j^{(2)}}))& \\ \cdot \mathrm{cum}_2(d^{(T)}(-\lambda_{j^{(1)}}), d^{(T)}(\lambda_{j^{(3)}}))& \\ \cdot \mathrm{cum}_2(d^{(T)}(\lambda_{j^{(2)}}), d^{(T)}(\lambda_{j^{(3)}}))\}.&\end{aligned}$$

If one now replaces the $\mathrm{cum}_2(\cdot)$ terms by the representation (2.7) and multiplies, one of the terms in the resulting sum is of the form

$$\frac{1}{T^{12}} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} R^{(T)}(\lambda_{j^{(1)}}, -\lambda_{j^{(2)}}) R^{(T)}(-\lambda_{j^{(1)}}, \lambda_{j^{(3)}}) R^{(T)}(\lambda_{j^{(2)}}, -\lambda_{j^{(3)}}).$$

Thus, when $\mathrm{cum}_3(\cdot)$ is normalized by $(\sqrt{T^2})^3$, it is bounded in absolute value by a $O(1)$ term, unlike the 1 dimensional case.

The Brillinger 1 dimensional proof depends on the nice $O(T^{-1})$ bound and Fourier properties. As nice as the cumulant method is, it does not carry over well to the random field case. This is why one must resort to the Rosenblatt method of proof, as further discussed in Section 4.

3. Explicit third cumulant calculation

Let $j^{(k)} = (j_1^{(k)}, j_2^{(k)})$ and $\lambda_{j^{(k)}} = (\lambda_{j_1^{(k)}}, \lambda_{j_2^{(k)}})$. Without loss of generality, one may center the data at the sample mean so that $d^{(T)}(0,0) = 0$. Consider the normalized third order cumulant of a sum of periodograms.

$$\begin{aligned} T^3 \operatorname{cum}_3 \Bigg\{ & \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(1)}=1}^{T-1} I^{(T)}(\lambda_{j^{(1)}}), \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(2)}=1}^{T-1} I^{(T)}(\lambda_{j^{(2)}}), \\ & \qquad \left(\frac{2\pi}{T}\right)^2 \sum_{j^{(3)}=1}^{T-1} I^{(T)}(\lambda_{j^{(3)}}) \Bigg\} \\ = \frac{1}{T^9} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} & \operatorname{cum}_3\{ d^{(T)}(\lambda_{j^{(1)}}) d^{(T)}(-\lambda_{j^{(1)}}), \\ & d^{(T)}(\lambda_{j^{(2)}}) d^{(T)}(-\lambda_{j^{(2)}}), \\ & d^{(T)}(\lambda_{j^{(3)}}) d^{(T)}(-\lambda_{j^{(3)}}) \}. \end{aligned}$$

Consider a specific indecomposable partition of three sets, that is each partition element is a set of size 2.

$$\begin{aligned} \frac{1}{T^9} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} & \{ \operatorname{cum}_2(d^{(T)}(\lambda_{j^{(1)}}), d^{(T)}(-\lambda_{j^{(2)}})) \\ & \times \operatorname{cum}_2(d^{(T)}(\lambda_{j^{(3)}}), d^{(T)}(-\lambda_{j^{(1)}})) \\ & \times \operatorname{cum}_2(d^{(T)}(\lambda_{j^{(2)}}), d^{(T)}(\lambda_{j^{(3)}})) \}. \end{aligned} \tag{3.1}$$

Examine the first second order cumulant or covariance in (3.1). Let $t = (t_1, t_2)$ and $u = (u_1, u_2)$.

$$\begin{aligned} & \operatorname{cum}_2(d^{(T)}(\lambda_{j^{(1)}}), d^{(T)}(-\lambda_{j^{(2)}})) \\ & = \sum_{t_1=0}^{T-1} \sum_{t_2=0}^{T-1} \sum_{u_1=-t_1}^{T-t_1-1} \sum_{u_2=-t_2}^{T-t_2-1} e^{-i\langle \lambda_{j^{(1)}}, u\rangle} e^{-i\langle t, (\lambda_{j^{(1)}} - \lambda_{j^{(2)}})\rangle} \\ & \quad \cdot c_2(u_1, u_2) I(|u_1| \le 1) I(|u_2| \le 1) \\ & = \sum_{u_1=-1}^{1} \sum_{u_2=-1}^{1} e^{-i\langle \lambda_{j^{(1)}}, u\rangle} c_2(u_1, u_2) \\ & \quad \cdot \Bigg(\sum_{t_1=0}^{T-1} \sum_{t_2=0}^{T-1} e^{-i\langle t, (\lambda_{j^{(1)}} - \lambda_{j^{(2)}})\rangle} \\ & \qquad \cdot I(-u_1 \le t_1 \le T - u_1 - 1) I(-u_2 \le t_2 \le T - u_2 - 1) \Bigg) \\ & = e^{-i\langle \lambda_{j^{(1)}}, (-1,-1)\rangle} c_2(-1, -1) \end{aligned}$$

$$
\begin{aligned}
&\cdot (\Delta^{(T)}(\lambda_{j_1^{(1)}} - \lambda_{j_1^{(2)}}) - 1)(\Delta^{(T)}(\lambda_{j_2^{(1)}} - \lambda_{j_2^{(2)}}) - 1) \\
&+ e^{-i\langle \lambda_{j^{(1)}}, (-1,0)\rangle} c_2(-1,0)(\Delta^{(T)}(\lambda_{j_1^{(1)}} - \lambda_{j_1^{(2)}}) - 1)\Delta^{(T)}(\lambda_{j_2^{(1)}} - \lambda_{j_2^{(2)}}) \\
&+ \\
&\vdots \\
&+ e^{-i\langle \lambda_{j^{(1)}}, (1,1)\rangle} c_2(1,1)(\Delta^{(T)}(\lambda_{j_1^{(1)}} - \lambda_{j_1^{(2)}}) - e^{-i(T-1)(\lambda_{j_1^{(1)}} - \lambda_{j_1^{(2)}})}) \\
&\cdot (\Delta^{(T)}(\lambda_{j_2^{(1)}} - \lambda_{j_2^{(2)}}) - e^{-i(T-1)(\lambda_{j_2^{(1)}} - \lambda_{j_2^{(2)}})}).
\end{aligned}
$$

The above is a sum of nine terms, labelled as $(-1,-1), \ldots, (1,1)$ say. Recall that we consider

$$\lambda_{j^{(k)}} = \left(\frac{2\pi}{T} j_1^{(k)}, \frac{2\pi}{T} j_2^{(k)}\right).$$

The product of the covariances in (3.1) involves 9^3 terms for each index of the six-fold sum. Consider the $(-1,-1)$ term above under multiplication of the covariances in equation (3.1). The resulting sum is of the form,

$$
\begin{aligned}
(3.2)\qquad \frac{1}{T^9} \sum_{j^{(1)}=1}^{T-1} \sum_{j^{(2)}=1}^{T-1} \sum_{j^{(3)}=1}^{T-1} \Bigg\{ &\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(1)} - j_1^{(2)})\right) - 1\right) \\
&\cdot \left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_2^{(1)} - j_2^{(2)})\right) - 1\right) \\
&\cdot \left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(3)} - j_1^{(1)})\right) - 1\right) \\
&\cdot \left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_2^{(3)} - j_2^{(1)})\right) - 1\right) \\
&\cdot \left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(2)} + j_1^{(3)})\right) - 1\right) \\
&\cdot \left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_2^{(2)} + j_2^{(3)})\right) - 1\right) \\
&\cdot e^{i(2\pi/T)j_1^{(2)}} e^{i(2\pi/T)j_2^{(2)}} (c_2(-1,-1))^3 \Bigg\}.
\end{aligned}
$$

After some careful algebra, it can be shown that the above sum is $O(T^{-3})$. A similar careful analysis can be employed to show that the sum of the product of the $(-1,-1)$ terms is $O(T^{-3})$. Each of the remaining 9^3 terms is likewise $O(T^{-3})$. Refer to the Appendix for further computational details.

4. A modification of Rosenblatt's result

This section will study the proper centering of sums of periodograms. This centering is at the expected value of the periodogram, which is a biased estimate of the spectral density. An alternative is to replace the definition of the periodogram with unbiased estimates of the spectral density, as suggested by Guyon (1982). In the one dimensional time series one can use the biased sample covariances, as for example in Brockwell and Davis (1991). The observation or comment being made here is to note that one needs to be careful about the centering in Z^2. This problem does not occur in a real time series, that is a process indexed by Z. This is discussed more at the end of this section. This problem was noted in Benn (1996) and Benn and Kulperger (1997), where a spatial data analysis was studied as an application of sums of periodograms in a spatial setting.

Rosenblatt (1985) studies sums of periodograms for stationary processes indexed by Z^d. The outline of his procedure is now given. His Theorem 5, p. 116 is a central limit theorem (CLT) for a mean 0, stationary process with finite $2+\delta$ moments. The observation is on the discrete lattice of points in the rectangle $[0, T-1]^d$. The normalizing rate is $T^{d/2}$. This result is then applied to the quadratic process, for given u, say $Y^{(u)}$ built up from $Y_{u,j} = X_j X_{j+u}$, which has expectation $c(u) = \mathrm{Cov}(X_0, X_u)$. Taking $\delta = 2$, that is Y having 4 moments, or equivalently X having 8 moments, gives his Theorem 6. Finally he applies this to sums of periodograms, which he rewrites as sums of empirical covariances; see Theorem 4.1 below.

At this point take $d = 2$, and examine how Theorem 5, p. 116 of Rosenblatt (1985), applies to a random field observed on $A_T = [0, T-1]^2$.

4.1 *Proofs*

The sample covariance at lag u, as comes into play in the sums of periodograms result, is defined by

$$\hat{c}(u) = \frac{1}{T^2} \sum_{t=0}^{T-|u|-1} X_t X_{t+u} = \frac{1}{T^2} \sum_{t=0}^{T-|u|-1} Y_{u,t}, \tag{4.1}$$

where the sum is a shorthand for the sum over $t = (t_1, t_2)$, $0 \le t_i \le T - |u_i| - 1$. Subtracting $\bar{X}$ is not important in the discussion below (due to the properties of $\Delta^{(T)}(\lambda)$), and so the sample covariances are computed in terms of the X data centered at $\mathrm{E}(X) = 0$. Also note that if u has negative components, (4.1) gives the usual formula by a translation of the index of summation.

Remark. For a given u, the divisor in (4.1) could be changed to $(T-|u_1|)(T-|u_2|)$ resulting in an unbiased estimator of $c(u)$. This idea was used in Guyon (1982). However this does not apply directly to the sums of periodograms, where one needs the same divisor for all $\hat{c}(u)$ in the Rosenblatt proof. Specifically the right hand side of (4.2) with $\hat{c}$ replaced by the unbiased estimators is no longer the squared modulus of the finite Fourier transforms. Alternatively one could correct for this, but the same bias problem as described below comes into play at that stage. Again we note that this method can make use of fast Fourier transforms in the computations, whereas Guyon's (1982) does not.

The d dimensional periodogram is

$$I^{(T)}(\lambda) = \frac{1}{(2\pi T)^d} d^{(T)}(\lambda) d^{(T)}(-\lambda) \tag{4.2}$$

$$= \frac{1}{(2\pi)^d} \sum_{|u| \leq T-1} \hat{c}(u) e^{-i\langle \lambda, u \rangle}.$$

This is obtained by directly substituting the definition of the finite Fourier transform and doing a change of variables.

The result of interest applies to the smoothed periodograms

$$\int_0^{2\pi} I^{(T)}(\lambda) A(\lambda) d\lambda$$

with a smooth weight function A. f_4 is the fourth order spectral density. Note the integral is over $\lambda \in (0, 2\pi]^d$. The weight function A is both integrable and square integrable. These types of integrals and sums were also of interest to Whittle to obtain estimates in stationary time series, obtaining what he called Gaussian estimates; see Whittle (1962). They have also been used in point processes by Brillinger (1976).

To deal with the convergence of the mean of the smoothed periodogram we also make the following damping assumption on the lag covariance function $c(u)$.

C1: $\sum_u |u_j||c(u)| < \infty$

THEOREM 4.1. (Rosenblatt (1985), Theorem 7, p. 118) *Let $\{X(t)\}$ be an ergodic strictly stationary random field that satisfies the assumptions of Theorem* A, *Rosenblatt* (1985). *Consider*

$$\int_0^{2\pi} I^{(T)}(\lambda) A(\lambda) d\lambda$$

which is a quadratic form in $X(t)$ with real-valued weight functions $A(\lambda)$ square integrable. The smoothed periodogram is asymptotically normal with mean

$$\int_0^{2\pi} f(\lambda) A(\lambda) d\lambda$$

and limiting covariance

$$(2\pi)^d \left\{ 2 \int_0^{2\pi} A(\lambda) A(\lambda) f^2(\lambda) d\lambda + \int_0^{2\pi} \int_0^{2\pi} f_4(\lambda, -\mu, \mu) A(\lambda) A(\mu) d\lambda d\mu \right\}.$$

From (4.2)

$$(2\pi)^d \int_0^{2\pi} I^{(T)}(\lambda) A(\lambda) d\lambda = \sum_{|u| \leq T-1} \hat{c}(u) a(u)$$

where $a(\cdot)$ are the Fourier coefficients of A, that is

$$a(u) = \frac{1}{(2\pi)^2} \int_0^{2\pi} A(\lambda) e^{-\langle \lambda, u \rangle} d\lambda.$$

Therefore

$$(2\pi)^d T^{d/2}\left\{\int_0^{2\pi} I^{(T)}(\lambda)A(\lambda)d\lambda - \int_0^{2\pi} f(\lambda)A(\lambda)d\lambda\right\}$$
$$= \sum_{|u|\le T-1} T^{d/2}\{\hat{c}(u)-c(u)\}a(u) + T^{d/2}\sum_{u\notin[-T+1,T-1]^d} c(u)a(u).$$

Under the conditions A integrable and the damping condition C1 on $c(u)$ (given just before Theorem 4.1), it follows that the second term on the right hand side tends to 0 as $T\to\infty$. The first term can be written as

$$\sum_{|u|\le T-1} T^{d/2}\{\hat{c}(u)-c(u)\}a(u) = \sum_{|u|\le L} T^{d/2}\{\hat{c}(u)-c(u)\}a(u) + R^{(T)}(L)$$

where $R^{(T)}(L)$ is the remainder from the normalized sum.

For $d=2$

$$R^{(T)}(L) = \sum_u I(L+1\le u_1\le T-1, |u_2|\le T-1)T(\hat{c}(u)-c(u))a(u)$$
$$= \sum_u I(-T+1\le u_1\le -L-1, |u_2|\le T-1)T(\hat{c}(u)-c(u))a(u)$$
$$= \sum_u I(L+1\le u_2\le T-1, |u_1|\le T-1)T(\hat{c}(u)-c(u))a(u)$$
$$= \sum_u I(-T+1\le u_2\le -L-1, |u_1|\le T-1)T(\hat{c}(u)-c(u))a(u).$$

The above terms are handled in a similar fashion, and so consider only the first term. It can be rewritten as

$$\sum_u I(L+1\le u_1\le T-1, |u_2|\le T-1)T(\hat{c}(u)-\mathrm{E}(\hat{c}(u))+\mathrm{E}(\hat{c}(u))-c(u))a(u).$$

The first part is handled directly from Theorem 5, p. 116, Rosenblatt (1985). The second part has a zero limit.

Thus by Theorem 5, Rosenblatt (1985) for $d=2$

$$T\left(\int_0^{2\pi} I^{(T)}(\lambda)A(\lambda)d\lambda \quad \int_0^{2\pi} f(\lambda)A(\lambda)d\lambda\right) \Rightarrow N(\mu,\sigma^2)$$

where $\Rightarrow$ means convergence in distribution. The limit mean μ and variance σ^2 of this limiting normal distribution must still be calculated. Since the random variables in question have several finite moments, the mean of the limit distribution equals the limit of the means.

Let $f^{(T)}(\lambda)=\mathrm{E}(I^{(T)}(\lambda))$. Then

$$f^{(T)}(\lambda) = \frac{1}{(2\pi)^2}\sum_{|u|\le T-1}\frac{(T-|u_1|)(T-|u_2|)}{T^2}c(u)e^{-i\langle\lambda,u\rangle}$$
$$= \frac{1}{(2\pi)^2}\sum_{|u|\le T-1}\left(1-\frac{|u_1|+|u_2|}{T}+\frac{|u_1u_2|}{T^2}\right)c(u)e^{-i\langle\lambda,u\rangle}.$$

Then

$$f^{(T)}(\lambda) - f(\lambda) = \frac{1}{(2\pi)^2} \sum_{|u| \le T-1} \left(-\frac{|u_1| + |u_2|}{T} + \frac{|u_1 u_2|}{T^2} \right) c(u) e^{-i\langle \lambda, u \rangle} + o(T^{-1})$$
$$- o(1).$$

However under the mixing condition,

$$\sum_u |u_j|^2 |c(u)| < \infty$$

one obtains

$$\lim_{T \to \infty} T \int_0^{2\pi} (f^{(T)}(\lambda) - f(\lambda)) A(\lambda) d\lambda = -\sum_u (|u_1| + |u_2|) c(u) a(u). \tag{4.3}$$

Thus

$$(2\pi)^d T^{d/2} \left\{ \int_0^{2\pi} I^{(T)}(\lambda) A(\lambda) d\lambda - \int_0^{2\pi} f(\lambda) A(\lambda) d\lambda \right\} \Rightarrow N(\mu, \sigma^2) \tag{4.4}$$

where

$$\mu = -\sum_u (|u_1| + |u_2|) c(u) a(u).$$

In particular using the limit distribution (4.4) would lead to an incorrect confidence interval, that is the confidence interval would not be centered properly.

How does one obtain a corresponding CLT with limiting mean zero? The easiest, and most natural way is to consider the centered integral

$$(2\pi)^d T^{d/2} \left\{ \int_0^{2\pi} I^{(T)}(\lambda) A(\lambda) d\lambda - \int_0^{2\pi} f^{(T)}(\lambda) A(\lambda) d\lambda \right\}$$

which then has no asymptotic bias with which to be concerned.

Why is this result different than a 1D time series, in which the centering at f or $f^{(T)}$ does not matter. The analogue of the bias (4.3) in 1D is

$$-\sqrt{T} \sum_u \frac{(|u_1| + |u_2|)}{T} c(u) a(u) = O(T^{-1/2}) \to 0.$$

This is very much related to standard computational simplifications made in time series; see for example Brockwell and Davis (1991), where they can interchangeably use divisors T and $T - |u|$ to work with biased or unbiased sample covariances. Note also that in the random field case, one cannot so easily change divisors for convenience, that is one has to decide at the outset to work properly centered variables.

5. Conclusions

This paper deals with some practical aspects of the time series methods in stationary random fields on a lattice. There are differences with respect to a time series in 1 dimension. A problem in studying smooth sums of periodograms is considered. In 1 dimension Brillinger (1981) has a very elegant proof of convergence using Fourier properties and associated asymptotics of the cumulants of finite Fourier transforms of the data series. We have shown that this does not carry over directly to a random field case, that is dimension 2. Thus, one can use Rosenblatt's method of proof (Rosenblatt (1985)). With this proof, one must be careful about the centering terms. If not one may end up with a limit normal distribution with non-zero mean, and hence incorrect confidence intervals. This problem was first noted in Guyon (1982), where he used unbiased covariance estimators. Our solution is to center the periodograms, as this allows one to retain the use of fast Fourier transforms.

Both the problems noted here are essentially geometric problems. This is because the size of the omitted terms or bias terms in the random field case ($d = 2$) is of the same size as the normalizing factor in the asymptotic normal result. In the case of a one dimensional time series, the bias terms are of a smaller order than the normalizing rate, and hence this problem does not occur.

Acknowledgements

The authors thank the referee for a careful reading of the paper and for helpful suggestions.

Appendix

Rewrite equation (3.2) as

$$\sum_D O(T^2)\Bigg\{\sum_{j_1^{(3)}=1}^{T-1}\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(3)}-j_1^{(1)})\right)-1\right)\tag{A.1}$$

$$\cdot\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(2)}+j_1^{(3)})\right)-1\right)$$

$$\cdot\sum_{j_2^{(3)}=1}^{T-1}\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_2^{(3)}-j_2^{(1)})\right)-1\right)$$

$$\cdot\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_2^{(2)}+j_2^{(3)})\right)-1\right)\Bigg\}$$

$$+\sum_{D^c}\Bigg\{\sum_{j_1^{(3)}=1}^{T-1}\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(3)}-j_1^{(1)})\right)-1\right)\tag{A.2}$$

$$\cdot\left(\Delta^{(T)}\left(\frac{2\pi}{T}(j_1^{(2)}+j_1^{(3)})\right)-1\right)$$

$$\cdot \sum_{j_2^{(3)}=1}^{T-1} \left(\Delta^{(T)} \left(\frac{2\pi}{T} (j_2^{(3)} - j_2^{(1)}) \right) - 1 \right)$$
$$\cdot \left(\Delta^{(T)} \left(\frac{2\pi}{T} (j_2^{(2)} + j_2^{(3)}) \right) - 1 \right) \Bigg\}$$

where D is the set of diagonal terms over the outer 4 sums of $j_1^{(1)}$, $j_2^{(1)}$, $j_1^{(2)}$, $j_2^{(2)}$ with

$$D = \{j_1^{(1)} = j_1^{(2)}, j_2^{(1)} = j_2^{(2)}\}.$$

In (A.1), the outer sum is over $(T-1)^2$ terms. The inner sum can now be obtained in $O(T)$'s as follows. For each $(j_2^{(1)}, j_2^{(2)})$ pair in D

$$\sum_{j_2^{(3)}=1}^{T-1} \left(\Delta^{(T)} \left(\frac{2\pi}{T} (j_2^{(3)} - j_2^{(1)}) \right) - 1 \right) \left(\Delta^{(T)} \left(\frac{2\pi}{T} (j_2^{(2)} + j_2^{(3)}) \right) - 1 \right)$$
$$= (T-1)(-1) + (T-3)(-1)^2 + (-1)(T-1)$$
$$= O(T).$$

The same argument shows the other inner sum over $j_1^{(3)}$ is $O(T)$. Thus equation (A.1) is $O(T^6)$ as there are $(T-1)^2 O(T) O(T)$ such terms.

Now re-write the set D^c as follows. Let

$$D_2\{(j_1^{(1)}, j_2^{(1)}, j_1^{(2)}, j_2^{(2)}) : j_1^{(1)} + j_1^{(2)} = T, j_2^{(1)} + j_2^{(2)} = T\}.$$

Then

$$D^c = D^c \cap (D_2 \cup D_2^c)$$
$$= (D^c \cap D_2) \cup (D^c \cap D_2^c).$$

Note $|D^c \cap D_2| = O(T^2)$.

Equation (A.2) can then be re-written as

$$(-1)^2 \sum_{D^c \cap D_2} (\cdot) + (-1)^2 \sum_{D^c \cap D_2^c} (\cdot).$$

For each element of $D^c \cap D_2$, the inner sums of equation (A.2) evaluate to

$$(T \quad 1)^2 \mid (T-2)(-1)^2 - O(T^2).$$

Thus, the first sum above is $O(T^6)$. On $D^c \cap D_2^c$, each inner sum of equation (A.2) evaluates to

$$(T-1)(-1) + (T-2)(-1)^2 = O(T).$$

Taking note that $|D^c \cap D_2^c| = O(T^4)$ shows that the second sum above is $O(T^6)$ also.

References

Benn, A. (1996). A Poisson driven stationary process model in spectroscopy, Ph.D. Thesis, Department of Statistical and Actuarial Sciences, University of Western Ontario, Canada.

Benn, A. and Kulperger, R. J. (1997). Integrated marked Poisson processes with application to image correlation spectroscopy, *Canad. J. Statist.*, **25**, 215–231.

Brillinger, D. R. (1976). Statistical inference for stationary point processes, *Stochastic Processes and Related Topics*, Volume 1 (ed. M. L. Puri), 55–99, Academic Press, New York.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, expanded ed., Holden-Day, San Francisco.

Brockwell, Peter J. and Davis, Richard A. (1991). *Time Series: Theory and Methods*, 2nd ed., Springer, New York.

Guyon, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice, *Biometrika*, **69**, 95–105.

Leonov, V. P. and Shiryaev, A. N. (1959). On a method of calculation of semi-invariants, *Theory Probab. Appl.*, **IV**, 319–329.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*, Birkhäuser, Basel.

Whittle, P. (1962). Gaussian estimation in stationary time series, *Bull. Inst. Internat. Statist.*, **39**, 105–129.